

## **¿Vivimos en Matrix? Entrevista a Donald Hoffman**

Básicamente la idea de la Teoría de la Interface es que no vemos el mundo como es, sino como conviene a nuestra especie. La realidad que percibimos no es la verdadera realidad sino que lo que percibimos es una especie de escritorio de Windows donde interactuamos con los objetos de una manera conveniente para nosotros. El sol , las montañas y los planetas son iconos en ese escritorio. Pero la verdadera realidad no la conocemos, está detrás y sería el equivalente al lenguaje de maquina y al hardware y a todos los componentes internos del ordenador.

Pero comentábamos también que Hoffman ha estirado esta teoría hasta desarrollar una teoría más completa, que no se refiere solo a la percepción sino que intenta responder nada menos que al problema de qué está hecho el Universo. Llama a esta teoría realismo consciente. Este realismo consciente combina elementos filosóficos con otros procedentes del mundo de la Física y las matemáticas de una manera poco convencional.

El punto de partida del realismo consciente es la conciencia. Es una solución monista (no dualista) en la que el mundo está compuesto por completo de agentes conscientes. Las interacciones entre estos agentes conscientes es lo que da lugar al mundo físico, a todos los objetos y propiedades que nosotros observamos. Hoffman propone de esta manera una solución al problema Mente-Cuerpo. Según él, explicar la conciencia a partir de la materia es muy difícil (nadie lo ha conseguido hasta la fecha) pero explicar matemáticamente la materia a partir de la conciencia dice que no lo es tanto. Y esa es la vía que ha tomado y está explorando en la actualidad, explicar cómo los agentes conscientes construyen los objetos físicos y sus propiedades. Para ello utiliza en su trabajo la teoría de juegos evolucionista y algoritmos genéticos.

El realismo consciente afirma que el cerebro no tiene poderes causales, el cerebro no causa la experiencia consciente. En cambio, ciertos agentes conscientes, cuando son activados por la interacción con otros agentes conscientes construyen los cerebros y el resto de nuestra anatomía como iconos de sus escritorios. Los cerebros no crean la conciencia : es la conciencia la que crea el cerebro. Como veis, son afirmaciones espectaculares y cuando intentamos avanzar en ellas nos encontramos con complicadas fórmulas matemáticas.

Por ello, no nos hemos podido resistir a entrevistar a Hoffman e intentar que nos explique con más detalle sus ideas en el lenguaje de la calle. Tú juzgarás si lo hemos conseguido pero creo que puedes disfrutar de una entrevista alucinante donde desgrana unas ideas que no escuchamos todos los días. Por otra parte, estamos ante un científico de una humanidad y amabilidad fuera de lo común que nos ha tratado con una sencillez y cordialidad asombrosas.

***1- Supongo que cuando da conferencias acerca de la teoría de la interface de la percepción la gente pensará en Matrix y le preguntará por ello. Yo voy a hacer lo mismo: ¿vivimos en Matrix? Vivimos dentro de una simulación de ordenador?***

La Teoría de la Interface de la Percepción (de ahora en adelante TIP) es similar a algunas ideas de The Matrix y radicalmente diferente de otras ideas de The Matrix

TIP es similar a Matrix en que dice que nuestras percepciones no revelan la realidad como es cuando no la estamos observando (de aquí en adelante diré “realidad” en lugar de la frase más larga: “realidad como es cuando no la estamos observando”). TIP dice que nuestras percepciones ocultan la realidad. De forma similar, en The Matrix los que están todavía enchufados a la matriz no ven la realidad, ven algo más placentero que la realidad.

Pero la TIP es diferente de Matrix en dos cosas importantes. Primero, la TIP propone sobre bases evolucionistas que los predicados de nuestras percepciones, tales como el espacio, el tiempo y los objetos físicos, son los predicados erróneos para describir la realidad. No hay nada en la realidad que sea en absoluto como el espacio, el tiempo y los objetos físicos. Por contra, en Matrix el mundo falso percibido por los que están enchufados a la matriz es un mundo que parece consistir en objetos físicos en el espacio y el tiempo; y cuando alguien, como Neo, se desconecta de la matriz y entra en el mundo real, ese mundo real consiste también en objetos físicos en el espacio y en el tiempo. En Matrix la única diferencia entre el mundo de la matriz y el mundo real son los objetos específicos que hay en cada mundo y lo que pasa con ellos. Por tanto, La TIP es más radical que Matrix. La TIP dice que si te pudieras desenchufar de tus percepciones, el mundo que encontrarías es totalmente diferente del mundo de tus percepciones, y casi seguro que no tiene tiempo y espacio.

La TIP es diferente de Matrix en un segundo aspecto. En Matrix el mundo falso de dentro de la matriz ha sido deliberadamente programado para engañar a los que son conectados a la matriz. Pero la TIP no requiere que nuestras percepciones hayan sido deliberadamente

programadas para engañarnos. Por contra, según la TIP nuestras percepciones no son como la realidad por dos razones. Primero, la realidad es infinitamente compleja y cada uno de nosotros tiene solamente unas capacidades perceptivas finitas. Así que nuestras percepciones omiten necesariamente casi toda la realidad. Segundo, las pobres capacidades perceptivas que tenemos son moldeadas por la evolución para ajustarse a las funciones de adaptación o ajuste que son relevantes para nosotros en nuestro nicho. Esto significa que nuestras capacidades perceptivas se usan para mantenernos con vida el tiempo necesario para reproducirnos. Sencillamente no podemos permitirnos el lujo de unas percepciones que describan la verdadera naturaleza de la realidad. Eso no es lo que necesitamos. La percepción no va de ver la realidad sino de tener hijos.

***2- Veo conexiones entre la TIP y la teoría de Jakob von Uexküll del Umwelt. Cada especie percibe el nicho del mundo en que se encuentra de la forma más conveniente para ella. La Umwelt sería el escritorio y el Umgebung sería la realidad tal como es (si es que existe tal cosa) ¿Está de acuerdo?***

Estoy de acuerdo en que la noción de Umwelt de von Uexküll es bastante similar a la noción de escritorio de la TIP. No estoy tan seguro de que la noción de Umgebung se corresponda con la noción de la TIP de la realidad como es. Algunas interpretaciones de la noción de Umgebung dicen que es una parte del Umwelt, la parte compuesta de las características distalmente orientadas de un mundo externo espacio temporal. Esta interpretación del Umgebung lo convierte en totalmente diferente de la noción de la TIP de la realidad como es, que la TIP propone que seguramente no es una realidad espacio-temporal.

***3- ¿Qué piensa del solipsismo? ¿Existe un mundo ahí fuera? ¿Podemos conocerlo? Si a la evolución no le importa la verdad, sólo la supervivencia y el éxito reproductivo, entonces nuestras capacidades cognitivas no están diseñadas para la objetividad y la verdad...¿podemos eludir esta restricción colocada por la selección natural y hacer ciencia?***

La TIP no implica solipsismo -que es la afirmación de que la única cosa que existe soy yo y mis percepciones. La TIP permite que haya una realidad más allá de mi y de mis percepciones. La TIP no afirma conocer cómo es esa realidad; sólo afirma, sobre bases evolutivas, que la realidad es casi con seguridad distinta de mis percepciones en términos de objetos físicos en el espacio-tiempo ¿Podemos usar el método científico para entender cómo podría ser la realidad? Sí. No hay, por supuesto, garantía de que tendremos éxito. Pero, como indico más abajo en mi respuesta a la pregunta 4, tampoco hay seguridad de

que fracasaremos. Una teoría simple y concreta de la realidad -que es describible con los mismos predicados que usamos para describir las percepciones- es casi con seguridad falsa. No hay problema. Esto es progreso: estamos rechazando una teoría falsa que ocurre que nos es muy querida. Ahora podemos seguir adelante hacia otras teorías que podemos especificar con precisión matemática, hacer predicciones empíricas que pueden ser testadas experimentalmente y entonces reevaluar a la luz de los resultados de los experimentos.

***4- Al final de su charla TED, en las preguntas, dice que lo que es cierto de la percepción (que no es fidedigna) puede no serlo de las matemáticas o de la lógica. ¿Puede explicarlo un poco más? Si nos fijamos en lo mal que entendemos las probabilidades (el caso del problema de Monty Hall, por ejemplo) o el gran número de sesgos cognitivos que tenemos, parece que nuestra racionalidad no es mucho mejor que nuestra visión. Quizás nuestra razón no es muy buena porque diferentes adaptaciones entran en conflicto...¿Que opina?***

Nuestro análisis de los juegos evolucionistas nos lleva a concluir que no hay presiones selectivas para que veamos la realidad como es, y que genéricamente hay presiones selectiva para no ver la realidad como es. Pero la situación podría ser diferente para las matemáticas y la lógica. Podría haber tareas específicas que necesitamos realizar para mejorar nuestra fitness, que requieren una limitada , y quizás específica de dominio, capacidad para las matemáticas y la lógica. Por ejemplo, en nuestras interacciones sociales se da a veces esta lógica: Si tú quieres este beneficio, entonces tienes que realizar esa acción. Necesitamos por lo menos una limitada capacidad de valorar este condicional como un prerrequisito para para los intercambios sociales, de otra manera los aprovechados podrían minar el sistema social. También, podríamos necesitar una limitada capacidad matemática para entender beneficios para la fitness: dos manzanas dan aproximadamente el doble de beneficio que una manzana.

De nuevo, las presiones selectivas son justo para unas capacidades con las matemáticas y la lógica que satisfacen lo necesario, es decir, los suficiente para vencer a la competencia. Por lo tanto, deberíamos esperar que si unos atajos y unos sesgos funcionan en ciertos contextos, entonces no hay presiones adicionales para conseguir que la matemática o la lógica sea exactamente perfecta. De vez en cuando, los genes para esas capacidades podrían coincidir en una persona particular y obtendríamos una capacidad más avanzada en lógica y matemáticas. Pero esa capacidad sería, al menos al principio, un spandrel (enjuta), es decir, algo que no fue directamente seleccionado (los huesos son blancos pero no porque hayan sido seleccionados así por la selección natural, sino que la selección

natural seleccionó para huesos fuertes, y el calcio, que ocurre que es blanco, es una solución a este problema de fuerza)

***5- ¿Puede la TIP ser útil en Psiquiatría? ¿Podemos sacar conclusiones acerca de los delirios y las alucinaciones con esta herramienta? ¿Cómo es el escritorio de una persona con esquizofrenia?***

Creo que la TIP puede ser un aspecto importante de un nuevo campo que podríamos llamar “Psicoterapia evolucionista”, en el que una nueva generación de teoría de asesoramiento se basa firmemente en un entendimiento evolutivo de la naturaleza humana.

¿Por qué, por ejemplo, la gente tiene fobias a las arañas y ardillas, pero no a las armas y SUVs, cuando hoy en día más gente muere por armas y SUVs que por arañas y ardillas? La respuesta está en el ambiente en el que nuestros sistemas cognitivos y perceptuales evolucionaron, un ambiente en el que las arañas y roedores eran una amenaza real. La selección natural dio a estos peligros un lugar especial en nuestra interface perceptiva, y los cableó directamente en un sistema profundo de alarma en el sistema límbico del cerebro.

Creo que al final todos los trastornos psicológicos descritos en el DSM necesitarán ser entendidos en términos de psicología evolucionista. ¿Cuáles son los módulos cognitivos y perceptivos “normales” que han sido moldeados en la naturaleza humana por la evolución? ¿Cuáles son los conflictos normales en las actividades de estos módulos? ¿Cuáles son las variaciones genéticas normales que llegan a variaciones entre individuos en el funcionamiento de estos módulos? ¿Por qué existen variaciones? ¿Cuál podría ser la función adaptativa de estas variaciones? ¿Cuándo las variaciones son tan extremas que son consideradas por los demás como problemáticas?

La evolución ha cableado estrategias reproductivas en los hombres que son significativamente diferentes de las que ha cableado en las mujeres. Interferencias entre estas estrategias pueden ser una gran fuente de tensiones y dolor en las relaciones y el matrimonio. ¿Cómo puede una nueva generación de teorías de asesoramiento usar estos conocimientos evolutivos para ayudar de manera inteligente a las parejas a disfrutar de relaciones mutuamente satisfactorias?

Nuestro amplio abanico de emociones -amor, odio, ira, enamoramiento, asco, deseo de venganza, sentimientos de competitividad- son soluciones adaptativas a problemas que enfrentaron nuestros ancestros y que todavía enfrentamos nosotros. Para tener una psicoterapia apropiada para las emociones, debemos entender su naturaleza adaptativa, los patrones de sucesos que normalmente y adaptativamente las disparan, y los patrones que las disparan de manera desadaptativa en nuestro nicho actual.

***6- Usted dice en su brillante charla TED que las neuronas y cerebros no son causas, sólo símbolos...Desde el punto de vista del realismo consciente ¿qué puede decir del problema de la causalidad? ¿Existen causas? ¿Qué son? ¿Dónde están?***

Podríamos resumir el resultado de nuestras simulaciones evolutivas diciendo que “el espacio-tiempo está condenado”, es decir, el espacio-tiempo y los objetos físicos en el espacio-tiempo no son la naturaleza fundamental de la realidad. Muchos físicos, como David Gross, Ed Witten y Nima Arkani-Hamed están también diciendo que el espacio-tiempo está condenado y que la teoría de cuerdas y los modelos *amplituedro* de amplitudes de dispersión nos dicen que existe una realidad fundamental -actualmente desconocida- detrás del velo del espacio-tiempo.

Esto implica que nuestra creencia actual de que los objetos físicos y el espacio-tiempo tienen poderes causales es falsa. Podríamos descubrir que esto es debido a que nada se corresponde con nuestra noción de causa. O podría ser debido al hecho de que algo fuera del espacio-tiempo es la verdadera fuente de causalidad en el mundo.

Estoy explorando la posibilidad de que la conciencia pudiera ser lo fundamental, y que “agentes conscientes” pudieran ser la verdadera fuente de causalidad en el mundo. Quizás las elecciones de agentes conscientes son causas genuinas con efectos genuinos sobre las experiencias de otros agentes conscientes. Merece la pena explorar esta posibilidad para ver si conduce a predicciones comprobables. La noción de causalidad que estoy explorando no es una elección de libre albedrío completamente no restringido, sino una elección de libre albedrío limitada en parte por las elecciones de libre albedrío de otros agentes conscientes que se combinan de forma conjunta para constituir dicho agente.

Esta idea puede ser totalmente errónea. Pero está en línea con similares ideas que el físico Chris Fuchs ha propuesto dentro del marco del bayesianismo cuántico o “QBism”

***7- Muy relacionado con la pregunta anterior está el problema del libre albedrío...¿Qué dice el realismo consciente acerca de él? ¿Existe libre albedrío en el mundo descrito pro su teoría?***

El libre albedrío podría ser la verdadera fuente de causalidad en el mundo. Esto sería un resultado fascinante. Veremos. Esta es la dirección que estoy explorando actualmente.

***8- Usted siempre tiene en cuenta la evolución en su trabajo pero el realismo consciente no es muy darwiniano. Por lo que sabemos la conciencia es un recién llegado en la historia evolutiva del universo. Primero des la materia, después la biología y luego aparece la conciencia. Pero usted dice que es al revés: la conciencia es lo primero. Usted dice que el realismo consciente no es pansiquismo ni idealismo...¿Puede explicarlo?***

La teoría estándar de la evolución biológica asume una ontología fisicalista, en la que el espacio-tiempo y los objetos (tales como organismos, recursos, ADN) son parte de la realidad objetiva y tienen genuinos poderes causales. Sin embargo, en el corazón de la evolución hay un algoritmo, llamado “Darwinismo Universal”, que se basa en variación, selección y retención. Este algoritmo abstracto es capturado en las matemáticas de la teoría evolucionista del juego, que es el marco que nosotros usamos para descubrir que la selección natural no favorece las percepciones realistas. Por tanto, el darwinismo universal y la teoría de juegos evolucionista ofrecen la predicción clara de que la ontología fisicalista es incorrecta. Se necesita una ontología diferente. Una teoría de la evolución en esta nueva ontología debe satisfacer una importante condición límite: cuando esta teoría se proyecta hacia atrás hacia la interface espacio-tiempo del Homo sapiens debe darnos la teoría fisicalista estándar de la evolución biológica.

El realismo consciente propone que la ontología correcta es una en la que la conciencia es fundamental. Esta es una hipótesis científica que debe ser sujeta a los procesos científicos normales de hacer predicciones que sean comprobadas en experimentos cuidadosos.

El realismo consciente dice que la conciencia, y no los objetos físicos en el espacio-tiempo, como rocas, árboles o quarks, es fundamental. Pero el realismo consciente no dice que las rocas, árboles y quarks son ellos mismos conscientes. Simplemente dice que las rocas, árboles y quarks están entre los símbolos que los agentes conscientes emplean, entre las experiencias que los agentes conscientes disfrutan. Así que el realismo consciente es muy diferente del pansiquismo, que permite que los objetos físicos estén entre los

elementos fundamentales de la realidad, pero afirma que los objetos físicos son ellos mismos conscientes.

El realismo consciente difiere también del idealismo de Berkeley, que afirma *esse est percipi*, es decir, ser es ser percibido, y que Dios es el percibido que todo lo ve y que sus percepciones hacen que las cosas existan. El realismo consciente permite que los agentes conscientes existan incluso cuando no son percibidos (contrariamente al *esse est percipi*) y no requiere el Dios omnisciente de Berkeley.

***9- En su artículo “Conscious Realism and the Mind-Body problem” no puedo ver claramente al final qué es un agente consciente y cómo los agentes conscientes producen la materia y la realidad. ¿Podría explicar qué son y cuál es nuestra relación con ellos?***

Una teoría detallada de los agentes conscientes se presenta en el artículo “Objects of consciousness” que fue publicado en 2014 en la revista *Frontiers of Psychology*. Está disponible gratis online en esta URL:

*<http://journal.frontiersin.org/article/10.3389/fpsyg.2014.00577/full>*

En resumen, un agente consciente es una teoría matemática de la conciencia que incorpora la idea de que la percepción, la decisión y la acción son todas propiedades esenciales de la conciencia. Esta idea está capturada en el lenguaje matemático de la teoría de la medida, los núcleos Markovianos y, equivalentemente, del álgebra geométrica.

La idea acerca de la relación con la física es que la dinámica asintótica (es decir, a largo plazo) de las interacciones de los agentes conscientes es lo que se representa por partículas físicas en el espacio-tiempo. Esperamos demostrar que las simetrías en la estructura de la dinámica de los agentes conscientes lleva a invariancia superconformal y doble conformal. Esto nos permitiría hacer conexiones directas con la teoría cuántica supersimétrica y el trabajo reciente sobre amplituedros

***10- ¿Qué lugar tiene la moralidad en el realismo consciente? ¿Cuál es el significado de bien y mal, egoísmo, justicia y altruismo?***

La moralidad será un tema muy interesante de explorar. Nuestras nociones convencionales aquí tendrán que ceder el paso ante nuevas ideas igual que nuestras nociones de espacio-



tiempo y objetos físicos han tenido que rendirse. Podría resultar que la moralidad no es un absoluto que se aplica de forma universal. En su lugar podría ser un código específico de especie que varía de especie a especie

Por ejemplo, un pájaro llamado “piquero camanay” (*Sula nebouxii*) pone dos huevos. Cuando nacen una de las crías mata a la otra, bien picoteándola hasta la muerte o expulsándola al exterior donde muere bajo la acción de los elementos. La madre mira el espectáculo y no interviene, después de todo ella mató también a su hermano.

Este fratricidio (siblicidio) es, creemos, inmoral para el *H. sapiens*. Pero sobre qué bases podemos justificar que es también inmoral para el piquero camanay? Este fratricidio podría ser esencial para la supervivencia del piquero en su nicho, e impedir el fratricidio podría ser un tipo de genocidio (o más bien de especiecidio)

Cuando los leones matan para comer no esperan educadamente para asegurarse de que todos reciben una porción justa. En vez de eso, el más fuerte come primero y con viciosos gruñidos o violencia directa aleja a los otros. Solo si sobra algo obtendrá el más débil su bocado. Esta conducta podría ser considerada inmoral por humanos. Para los leones, sin embargo, puede ser esencial para la supervivencia de la especie.

Estas observaciones podrían informar una nueva teoría de la moralidad para los agentes conscientes en la que el código moral no es absoluto entre todos los agentes, sino que varía de maneras interesantes dependiendo del contexto y la constitución de los agentes. La moneda en esta esfera podría ser la información, y la moralidad podría tratar de las formas adecuadas en que hay que intercambiar la información.

### ***11- ¿En qué está trabajando ahora? ¿Qué misterio le gustaría desvelar?***

Estoy deseando conectar las matemáticas de los agentes conscientes con trabajos recientes con amplitud y simetría conformacional en física. Me gustaría desvelar el misterio de qué hay detrás del espacio-tiempo. Creo que los físicos han descubierto muchas pistas importantes en forma de varias simetrías que deben ser obedecidas por lo que sea que está detrás del espacio-tiempo. Quiero usar estas pistas para guiar el refinamiento detallado de la teoría de los agentes conscientes y sus interacciones. La posibilidad de esa síntesis es excitante. Los desafíos técnicos son sustanciales y me estoy esforzando en aprender las herramientas matemáticas y físicas necesarias.

Sería maravilloso si la teoría de los agentes conscientes produjera una nueva manera de pensar acerca de las relaciones entre ciencia y religión. Mi opinión es que puede haber una relación de sinergia entre ambas, pero requeriría que la religión renuncie a su dogmatismo y su insistencia en una autoridad no cuestionada de las sagradas escrituras, y requeriría que los científicos renuncien al espacio-tiempo y el fisicalismo. No es una tarea fácil para ambos lados. Pero ambos lados tiene algo profundo que decir al otro. Las religiones han afirmado desde hace mucho que los objetos físicos del espacio-tiempo no son la realidad última; la ciencia ha afirmado desde hace mucho que el dogmatismo y la autoridad deben rendirse a la lógica y le experimento. Una humanidad que pueda coordinar los conocimientos de ciencia y religión y soltar los defectos de cada una de ellas es una humanidad que caminará más inteligente y pacíficamente hacia el futuro.